Rworksheet_cadiz#3b

John Dave R. Cadiz

2024-10-02

1. Create a data frame using the table below..

```
## a. Write the codes.
data <- data.frame(
                  Respondents = 1:20,
                  Sex = c(2, 2, 1, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 1,
                          1, 2),
                  FathersOccupation = c(1, 3, 1, 3, 1, 2, 3, 1, 1, 1, 3, 2, 1,
                                        3, 1, 2, 3, 1, 2, 1),
                  PersonsAtHome = c(5, 7, 3, 8, 5, 9, 6, 7, 8, 4, 7, 5, 4, 7,
                                    8, 8, 3, 11, 7, 6),
                  SiblingsAtSchool = c(6, 4, 4, 1, 2, 3, 5, 3, 2, 4, 2, 5, 5,
                                       5, 2, 1, 2, 5, 3, 2),
                  TypesOfHouses = c(1, 2, 3, 1, 1, 3, 3, 1, 2, 3, 2, 3, 2, 2,
                                    3, 3, 3, 3, 3, 2))
```

```
## b. Describe the data. Get the structure or the summary of the data
#The data consist of 20 respondents together with the variable names such as sex,
#fathers occupation, Persons at home, Siblings at school, and the type of houses
str(data)
```

```
## 'data.frame':    20 obs. of  6 variables:
##  $ Respondents      : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Sex              : num  2 2 1 2 2 2 2 2 1 2 ...
##  $ FathersOccupation: num  1 3 1 3 1 2 3 1 1 1 ...
##  $ PersonsAtHome    : num  5 7 3 8 5 9 6 7 8 4 ...
##  $ SiblingsAtSchool : num  6 4 4 1 2 3 5 3 2 4 ...
##  $ TypesOfHouses    : num  1 2 3 1 1 3 3 1 2 3 ...
```

```
summary(data)
```

```
##   Respondents         Sex       FathersOccupation PersonsAtHome
##  Min.   : 1.00   Min.   :1.00   Min.   :1.0       Min.   : 3.0
##  1st Qu.: 5.75   1st Qu.:1.75   1st Qu.:1.0       1st Qu.: 5.0
##  Median :10.50   Median :2.00   Median :1.5       Median : 7.0
##  Mean   :10.50   Mean   :1.75   Mean   :1.8       Mean   : 6.4
##  3rd Qu.:15.25   3rd Qu.:2.00   3rd Qu.:3.0       3rd Qu.: 8.0
##  Max.   :20.00   Max.   :2.00   Max.   :3.0       Max.   :11.0
##  SiblingsAtSchool TypesOfHouses
##  Min.   :1.0      Min.   :1.0
##  1st Qu.:2.0      1st Qu.:2.0
##  Median :3.0      Median :2.5
##  Mean   :3.3      Mean   :2.3
##  3rd Qu.:5.0      3rd Qu.:3.0
```

```
##  Max.   :6.0      Max.   :3.0
```

```
## c. Is the mean number of siblings attending is 5?
meansib <- mean(data$SiblingsAtSchool)
meansib ==5
```

```
## [1] FALSE
```

```
## d. Extract the 1st two rows and then all the columns using the subsetting functions.
## Write the codes and its output.
firstTwoRows <- data[1:2, ]
print(firstTwoRows)
```

```
##   Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1           1   2                 1            5                6
## 2           2   2                 3            7                4
##   TypesOfHouses
## 1             1
## 2             2
```

```
## e. Extract 3rd and 5th row and 4th column. Write the code and its result.
exrow3n5col4 <- data[c(3,5), c(2,4)]
print(exrow3n5col4)
```

```
##   Sex PersonsAtHome
## 3   1             3
## 5   2             5
```

```
## f. Select the variable types of the houses then store the vector that result as
#type_houses. Write the codes

type_houses <- data$TypesOfHouses
print(type_houses)
```

```
##  [1] 1 2 3 1 1 3 3 1 2 3 2 3 2 2 3 3 3 3 3 2
```

```
## g. Select only all Males respondent that their father occupation was farmer.
## Write the codes and its output.

maleFarmer <- subset(data, Sex == 1 & FathersOccupation == 1)
print(maleFarmer)
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 3            3   1                 1            3                4
## 9            9   1                 1            8                2
## 18          18   1                 1           11                5
##    TypesOfHouses
## 3              3
## 9              2
## 18             3
```

```
## h. Select only all females respondent that have greater than or equal to 5 number
## of siblings attending school. Write the codes and its output.
femaleSibs <- subset(data, Sex == 2 & SiblingsAtSchool >= 5)
print(femaleSibs)
```

```
##    Respondents Sex FathersOccupation PersonsAtHome SiblingsAtSchool
## 1            1   2                 1            5                6
## 7            7   2                 3            6                5
```

```
## 12          12   2                 2             5                5
## 13          13   2                 1             4                5
##    TypesOfHouses
## 1              1
## 7              3
## 12             3
## 13             2
```

2. Write a R programe to create an empty data frame. Using the following codes:

```
df = data.frame(Ints=integer(),

Doubles=double(), Characters=character(),
Logicals=logical(),
Factors=factor(),
stringsAsFactors=FALSE)

print("Structure of the empty dataframe:")
```

```
## [1] "Structure of the empty dataframe:"
```

```
print(str(df))
```

```
## 'data.frame':    0 obs. of  5 variables:
##  $ Ints      : int
##  $ Doubles   : num
##  $ Characters: chr
##  $ Logicals  : logi
##  $ Factors   : Factor w/ 0 levels:
## NULL
```

```
## a. Describe the results.
## It shows that the data frame is empty and represents each different data types
## in five columns, since there is no data the observations remain 0.
```

3. Create a .csv file of this. save it as HouseholdData.csv

```
## a. Import the csv file into the R environment. Write the codes
data <- read.csv("HouseholdData.csv", header = TRUE, stringsAsFactors = FALSE)
print(data)
```

```
##    Respondents    Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   Male                 1              5                  2
## 2            2 Female                 2              7                  3
## 3            3 Female                 3              3                  0
## 4            4   Male                 3              8                  5
## 5            5   Male                 1              6                  2
## 6            6 Female                 2              4                  3
## 7            7 Female                 2              4                  1
## 8            8   Male                 3              2                  2
## 9            9 Female                 1             11                  6
## 10          10   Make                 3              6                  2
##    Types.of.Houses
## 1             Wood
```

```
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```
## b. Convert the Sex into factor using factor() function and change it into
## integer.[Legend: Male = 1 and Female = 2]. Write the R codes and its output.
data$Sex <- factor(data$Sex,
                   levels = c("Male", "Female"),
                   labels = c(1,2))
data$Sex <- as.integer(as.character(data$Sex))
print(data)
```

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   1                  1              5                  2
## 2            2   2                  2              7                  3
## 3            3   2                  3              3                  0
## 4            4   1                  3              8                  5
## 5            5   1                  1              6                  2
## 6            6   2                  2              4                  3
## 7            7   2                  2              4                  1
## 8            8   1                  3              2                  2
## 9            9   2                  1             11                  6
## 10          10  NA                  3              6                  2
##    Types.of.Houses
## 1             Wood
## 2         Congrete
## 3         Congrete
## 4             Wood
## 5    Semi-congrete
## 6    Semi-congrete
## 7             Wood
## 8    Semi-congrete
## 9    Semi-congrete
## 10        Congrete
```

```
## c. Convert the Type of Houses into factor and change it into integer. [Legend:
## Wood = 1; Congrete= 2; Semi-Congrete = 3]. Write the R codes and its output.
data$TypesOfHouses <- factor(data$Types.of.Houses,
                             levels = c("Wood", "Congrete", "Semi-Congrete"),
                             labels = c(1, 2, 3))
data$Types.of.Houses <- as.integer(data$Types.of.Houses)
```

```
## Warning: NAs introduced by coercion
```

```
print(data)
```

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   1                  1              5                  2
## 2            2   2                  2              7                  3
## 3            3   2                  3              3                  0
```

```
## 4            4   1                  3                8                  5
## 5            5   1                  1                6                  2
## 6            6   2                  2                4                  3
## 7            7   2                  2                4                  1
## 8            8   1                  3                2                  2
## 9            9   2                  1               11                  6
## 10          10  NA                  3                6                  2
##    Types.of.Houses TypesOfHouses
## 1               NA             1
## 2               NA             2
## 3               NA             2
## 4               NA             1
## 5               NA          <NA>
## 6               NA          <NA>
## 7               NA             1
## 8               NA          <NA>
## 9               NA          <NA>
## 10              NA             2
```

```
## d. On father's occupation, factor it as Farmer = 1, Driver = 2, and other = 3
## What is the R codes and its output.
data$Fathers.Occupation <- factor(data$Fathers.Occupation,
                                  levels = c(1, 2, 3),
                                  labels = c("Farmer", "Driver", "Others"))
data$Fathers.Occupation <- as.character(data$Fathers.Occupation)
print(data)
```

```
##    Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 1            1   1             Farmer               5                  2
## 2            2   2             Driver               7                  3
## 3            3   2             Others               3                  0
## 4            4   1             Others               8                  5
## 5            5   1             Farmer               6                  2
## 6            6   2             Driver               4                  3
## 7            7   2             Driver               4                  1
## 8            8   1             Others               2                  2
## 9            9   2             Farmer              11                  6
## 10          10  NA             Others               6                  2
##    Types.of.Houses TypesOfHouses
## 1               NA             1
## 2               NA             2
## 3               NA             2
## 4               NA             1
## 5               NA          <NA>
## 6               NA          <NA>
## 7               NA             1
## 8               NA          <NA>
## 9               NA          <NA>
## 10              NA             2
```

```
## e. Select only all females respondent that has a father whose occupation is driver.
## Write the R codes and its output.

femRes <- subset(data, Sex == 2 & Fathers.Occupation == "Driver")
print(femRes)
```

```
##   Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 2           2   2             Driver               7                  3
## 6           6   2             Driver               4                  3
## 7           7   2             Driver               4                  1
##   Types.of.Houses TypesOfHouses
## 2              NA             2
## 6              NA          <NA>
## 7              NA             1
```

```
## f. Select the respondents that have greater than or equal to 5 numbers of siblings
## attending school. Write the codes and its output.

sib5 <- subset(data, Siblings.at.School >= 5)
print(sib5)
```

```
##   Respondents Sex Fathers.Occupation Persons.at.Home Siblings.at.School
## 4           4   1             Others               8                  5
## 9           9   2             Farmer              11                  6
##   Types.of.Houses TypesOfHouses
## 4              NA             1
## 9              NA          <NA>
```